

## VALIDITAS ISI ALAT UKUR PENELITIAN: KONSEP DAN PANDUAN PENILAIANNYA

Helli Ihsan

Universitas Pendidikan Indonesia

Helli\_psi@upi.edu

### Abstrak

Ada banyak definisi tentang validitas isi tetapi Haynes dkk.(1995) berhasil membuat definisi yang sangat baik sehingga setiap frasa dari definisi itu diuraikan dengan baik pula olehnya. Dari uraian yang dia buat maka penulis berusaha mengaplikasikannya dalam bentuk pembuatan tabel penilaian validitas isi. Panduan penilaian itu berisikan penilaian kuantitatif terhadap elemen-elemen alat ukur, kerelevanan pernyataan terhadap definisi konseptual dan operasionalnya dan seberapa baik struktur bahasa dan kesederhanaan kalimat dalam menguraikan pernyataan. Peanduan ini bisa dijadikan pedoman pengujian validitas isi dalam pembuatan alat ukur di Indonesia.

### Definisi validitas isi

Validitas adalah sejauhmana tes itu mengukur apa yang dimaksudkan untuk diukur. Validitas alat ukur tidak terganggu. Secara umum ada tiga pendekatan dalam meneliti validitas suatu alat ukur, yaitu 1) validitas isi, 2) validitas konstruk, dan 3) validitas kriteria (Suryabrata, 2005). Validitas isi adalah validitas yang fokus kepada elemen-elemen apa yang ada dalam ukur (Coaley, 2010), sehingga analisis rasional adalah proses utama yang dilakukan dalam analisis validitas isi (Azwar, 2005). Validitas kriteria adalah mengaitkan alat ukur dengan alat ukur lain sebagai kriteria, apakah alat ukur itu bisa dijelaskan hasil korelasinya dengan dengan kriterianya berdasarkan teori yang ada (Devellis, 2003). Validitas konstruk adalah sebuah gambaran yang menunjukkan sejauhmana alat ukur itu menunjukkan hasil yang sesuai dengan teori (Azwar, 2005). Proses pengujian validitas konstruk adalah menghubungkan alat ukur itu dengan alat ukur lain yang memiliki kesamaan konsep atau dengan alat ukur-alat ukur lain yang secara teoritis berkaitan dengannya (Murphy & Davidshofer, 1991).

Secara umum definisi validitas isi adalah sejauhmana elemen-elemen instrumen asesmen relevan dan mewakili konstruk alat ukur yang ditargetkan untuk tujuan tertentu (Haynes, dkk. 1995). Validitas isi adalah penjelasan tentang suatu alat ukur secara substantif atau disebut validitas substantif yang fokus kepada konseptualisasi dan sejauhmana konsep-konsep sebelumnya yang ditampilkan dalam kajian literatur (Clark & Watson, 1995). Dari definisi yang dikemukakan oleh Haynes dkk (1995) dapat dijabarkan dalam berbagai istilah yang lebih luas.

“Instrumen asesmen” adalah alat yang digunakan untuk menerapkan semua metode validitas isi. “Elemen” dari instrumen asesmen adalah semua aspek proses pengukuran yang dapat mempengaruhi data yang diperoleh. Misalnya, elemen kuesioner seperti item pernyataan, format jawaban, dan instruksi. Elemen observasi perilaku seperti kode observasi, parameter sampling-waktu, dan situasi dimana observasi terjadi.

Kata “sejauhmana” berarti sejauhmana validitas isi telah dinilai atau memperoleh judgment secara kuantitatif (Haynes, Richard,

& Kubany, 1995). Hasil akhir dari validitas isi adalah penilaian tentang kelayakan isi tes. Lawshe mengajukan cara yang disebut rasio validitas isi sebagai hasil ukur penilaian ahli atas kesetujuannya terhadap validitas isi yang menggunakan teknik statistik kesepahaman antar penilai. Sejauhmana ini juga bisa ditunjukkan dengan seberapa detail penjelasan tentang alat ukur itu. Semakin detail maka akan semakin diyakini bahwa alat ukur itu semakin tepat (Murphy & Davidshofer, 1991).

Istilah “konstruk” berarti konsep, atribut, atau variabel yang menjadi target pengukuran. Frasa “tujuan tertentu” artinya bahwa kerelevanan dan keterwakilan instrumen asesmen tergantung kepada fungsi asesmen. Validitas isi adalah sebuah pernyataan bukan sebuah perilaku yang dicerminkan dari skor yang diperoleh instrumen asesmen. Validitas isi berbeda dengan skor kesimpulan yang diperoleh dari data asesmen. Misalnya

validitas isi kuesioner yang mengukur gejala depresi bisa tepat digunakan dalam fase screening tapi tidak tepat untuk fase perencanaan perlakuan (Mesick, 1993).

Kata “relevan” berarti ketepatan elemen pengukuran dengan target konstruk dan fungsi asesmen (Ebel, 1991). Misalnya, kerelevanan kuesioner report yang mengukur tingkatan gangguan stress pascatraumatik (PTSD) akan sangat berbeda dengan alat ukur yang mencakup item-item yang mengungkapkan tahapan PTSD seperti recurrent and distressing recollections and dreams, insomnia, dan hypervigilance. Kerelevanan akan mengurangi item-item yang tidak dalam domain PTSD (seperti item-item yang mencerminkan gejala ketergantungan obat atau gangguan konversi). Berikut adalah tabel yang bisa dijadikan pedoman dalam menentukan kerelevanan elemen-elemen suatu alat ukur (Haynes, Richard, & Kubany, 1995):

Elemen	Metode Asesmen			
	Kuesioner	Observasi Perilaku	Psikofisiologi	Self Monitoring
Mendeskripsikan item (Pernyataan, kode, skala jawaban)	R	R	R	R
Ketepatan penyusunan kalimat atau definisi setiap item	R	R	T	R
Bentuk jawaban item (mis. Skala jawaban)	R	T	T	R
Rangkaian item atau stimulus	R	T	R	T
Instruksi untuk partisipan	R	R	R	R
Waktu dalam menjawab	R	T	T	R
Sampel situasi	R <sup>a</sup>	R	R	R
Sampel perilaku atau peristiwa	R	R	R	R
Komponen jumlah, faktor, kelas jawaban	R	R	R	R
Metode dan standarisasi administrasi	R	R <sup>b</sup>	R	R
Penyekor, reduksi data, pembobotan item	R	R	R	R
Parameter sampling waktu	T	R	R	R
Menghadirkan stimulus	T	R	R	T
Definisi dan domain konstruk	R	R	R	R
Penyesuaian metode-modus	R	R	R	R
Penyesuaian fungsi-instrumen				

R = relevan, T = Tidak relevan

<sup>a</sup> Pada sebagian besar instrumen self-report, perlu untuk memperhatikan apakah jawaban merujuk kepada “saat ini”, “di waktu yang lalu”, dll. <sup>b</sup> Standarisasi perlu untuk antar partisipan atau antar alat asesmen, tergantung apakah pendekatan nomotetik atau idiografik digunakan, <sup>c</sup> pola nilai dan waktu dari pengukuran. <sup>d</sup> Misalnya, presentasi audio atau video sebagai stressor dalam asesmen laboratorium psikofisiologis; skenario observasi laboratorium.

“Representative atau keterwakilan” dari instrumen asesmen adalah sejauhmana elemen-elemen itu proporsional dalam mewakili bagian-bagian konstruk yang ditargetkan. Indikator-indikator perilaku yang dicantumkan sudahkah mencukupi dari aspek-aspek yang seharusnya ada. Item-item yang ditulis, apakah juga mewakili dari indicator yang dicantumkan.

Validitas isi ini sangat penting posisinya untuk pengujian validitas konstruk. Alat ukur yang tidak valid secara isi maka tidak perlu lagi diuji kevalidannya dengan metode validitas konstruk. Jadi alat ukur yang valid harus pernah diuji kevalidannya melalui uji validasi isi sebelum ke uji validasi yang lainnya. Validasi isi lebih menitikberatkan kepada validasi rasional atau logis sementara validasi yang lainnya ke validasi empiris. Validasi isi adalah dasar untuk pengujian validitas konstruk. Definisi validitas konstruk yaitu sejauhmana instrumen mengukur konstruk yang ingin dicapai, sangat membutuhkan proses yang dilakukan dalam validasi isi. Validasi konstruk meliputi semua validasi yang lainnya, yaitu validasi isi, konkuren, dan prediktif.

## **Metode Validitas Isi**

Untuk mewujudkan validitas isi yang baik harus digunakan dengan beragam metode, baik secara kualitatif atau kuantitatif, untuk menilai semua elemen instrumen asesmen. Pada tahap awal pengembangan instrumen tujuan validasi isi adalah mengurangi variasi potensi kesalahan pembuatan instrumen dan meningkatkan kemungkinan diperolehnya indeks validitas konstruk dalam studi lanjutan. Karena sumber-sumber kesalahan antara konstruk teoritis, metode asesmen, dan fungsi asesmen berbeda-beda, maka metode validasi isinya juga akan berbeda (Haynes, Richard, & Kubany, 1995).

Anjuran-anjuran dalam merancang validasi alat ukur dilihat dari sisi isi (Haynes, Richard, & Kubany, 1995), yaitu:

1. Berhati-hati dalam mendefinisikan domain dan aspek konstruk dan memperhatikan validasi isi sebelum mengembangkan elemen-elemen lain dari instrumen asesmen.

Fase ini adalah fase paling sulit dalam validitas isi. Pendefinisian sangat sulit dilakukan. Memahami definisi dimulai dari definisi konseptual yang jelas dan dapat dipahami dengan baik oleh penyusun alat ukur. Definisi konseptual yang kurang dipahami akan berakibat fatal dalam pembuatan definisi operasional dan berakibat pada kesalahan maksud dari apa yang ingin diukur. Maka dari itu sebelum melangkah ke langkah-langkah berikutnya seperti menyusun item-item, lebih baik kalau definisi konseptual dan operasionalnya dinilai atau dikaji oleh ahlinya.

2. Perhatikan semua elemen instrumen asesmen untuk melakukan validasi isi.

Elemen-elemen dari instrumen seperti format jawaban dan instruksi harus diperhatikan agar sesuai dengan tipe alat ukur. Instruksi “Jawablah pertanyaan-pertanyaan berikut dengan benar” tidak tepat untuk alat ukur subyektif seperti kuesioner dan self report, dan self monitoring. Variable-variabel seperti self esteem, konsep diri atau motivasi tidak tepat jika menggunakan instruksi seperti yang diatas. Instruksi diatas hanya tepat jika digunakan untuk tes-tes kemampuan maksimal seperti tes inteligensi dan bakat.

3. Gunakan populasi dan tahapan penyampelan untuk memperoleh item-item awal dan elemen-elemen lain.

Untuk mengetahui perilaku yang ingin diukur sebaiknya menggunakan banyak sumber informasi tentang perilaku itu. Selain para ahli yang memahami betul perilaku itu, mewawancarai subyek dari populasi yang menjadi target pengukuran perlu juga dilakukan untuk memperoleh gambaran perilaku yang lebih komprehensif, relevan dan representative terhadap aspek-aspek konstruk.

4. Gunakan penilaian dari penilai yang beragam untuk validasi isi dan kuantitatifkan penilaian yang menggunakan prosedur penskalaan yang formal.

Setiap elemen dalam instrumen harus dinilai apakah relevan atau tidak digunakan dalam pengukuran perilaku yang ingin diukur. Setiap elemen ini harus dinilai oleh beberapa ahli. Penilaian yang diberikan

itu menggunakan skala penilaian dengan lima atau tujuh. Penilaian itu ditujukan kepada relevansi, representasi, spesifikasi, dan kejelasan alat ukur. Statistic deskriptif terhadap penilaian itu bisa menjadi patokan untuk memutuskan tingkat kevalidan alat ukur itu.

5. Uji keterwakilan secara proporsional item-item yang digunakan.

Item-item dalam instrumen harus tersebar atau memiliki bobot yang relative di masing-masing aspek konstruk. Jangan sampai ada aspek yang memiliki item yang berlebihan dan aspek lain memiliki item yang kurang memadai. Kalau ada aspek yang secara rasional seharusnya memiliki item yang lebih banyak dari aspek lain maka aspek harus memiliki item yang banyak. Sekalipun tidak harus sama jumlah item dalam setiap aspek tetapi aspek-aspek atau dimensi dalam konstraknya harus proporsional dalam memiliki jumlah item.

6. Laporkan hasil validasi isi jika menerbitkan instrumen asesmen yang baru.

Indeks atau hasil secara kuantitatif terhadap validitas isi akan membantu masyarakat untuk mengetahui sejauhmana alat ukur telah melalui proses validasi ini. Angka-angka yang menunjukkan penilaian terhadap, misalnya kerelevanan, kerepresentatifan, sangat penting untuk member gambaran yang jelas terhadap pengguna alat ukur itu. Dengan indeks ini maka calon pengguna alat ukur akan merasa yakin bahwa alat ukur yang akan digunakannya berkualitas baik atau buruk.

**Panduan penilaian validitas isi**

Dibawah ini penulis berusaha membuat sebuah panduan penilaian atau judgment validitas isi yang menggunakan berbagai elemen alat ukur sebagai hal-hal yang dinilai oleh para ahli. Dengan panduan penilaian kuantitatif ini maka akan dengan mudah menilai seberapa baik validitas isi alat ukur. Dalam deskripsi alat ukur penyusunnya harus menguraikan elemen-elemen dalam alat ukur untuk kemudian dinilai oleh penilai. Mulai dari alat ukurnya sampai kepada waktu pengisian instrumen.

**DESKRIPSI ALAT UKUR**

NAMA ALAT UKUR :

1. Definisi Konseptual : (Definisi dalam teori)
2. Definisi Operasional : (Jumlah dan Definisi Indikator-Indikator)
3. Skala : (Skala Likert, Semantic Differential, Tes, atau lainnya)
4. Fungsi Instrumen : (Untuk seleksi, untuk self report, atau untuk self monitoring atau lainnya)
5. Instruksi : (Perintah untuk responden dalam menjawab)
6. Format Jawaban : (pola pilihan-pilihan jawaban yang diberikan)
7. Jumlah Item : (Jumlah pernyataan untuk setiap indikator atau satu alat ukur)
8. Penyekoran : (Cara penyekoran dalam menentukan pemberian asesmen)
9. Sampel populasi : (Jumlah dan karakteristik sampel responden yang akan diujicobakan)

10. Waktu : Interval waktu perilaku yang diukur dengan waktu pengukuran

**PENILAIAN VALIDASI ISI**

Berilah penilaian Anda terhadap aspek-aspek penilaian yang dimiliki oleh alat ukur yang kami susun ini. Penilaian yang anda berikan memiliki rentang 1 sampai 5. Semakin mendekati angka 1 maka penilaian anda semakin tidak relevan, memadai atau sesuai dan semakin mendekati 5 maka penilaian anda semakin relevan, memadai atau sesuai. Nilailah angka 3 jika penilaian anda berada di tengah-tengah.

**Tabel 1: Penilaian relevansi dan representasi aspek-aspek dalam instrumen**

No.	Aspek Penilaian	Penilaian					Masukan
1.	Definisi Konseptual	1	2	3	4	5	
2.	Definisi Operasional	1	2	3	4	5	
3.	Skala	1	2	3	4	5	
4.	Fungsi Instrumen	1	2	3	4	5	
5.	Instruksi untuk Responden	1	2	3	4	5	
6.	Representasi Jumlah Item	1	2	3	4	5	
7.	Format Jawaban	1	2	3	4	5	
8.	Penyekoran	1	2	3	4	5	
9.	Sampel populasi	1	2	3	4	5	
10.	Waktu	1	2	3	4	5	

**Rata-Rata Penilaian: .....**

**Tabel 2: Penilaian Ketepatan Tata Bahasa Kalimat**

No.	Indikator	Pernyataan	Penilaian					Masukan
			1	2	3	4	5	
1.		Pernyataan 1	1	2	3	4	5	
2.		Pernyataan 2	1	2	3	4	5	
3.		Pernyataan 3	1	2	3	4	5	

Rata-Rata Penilaian: .....

**Tabel 3: Kesesuaian pernyataan dengan definisi konsep dan operasional (teori)**

No.	Indikator	Pernyataan	Penilaian					Masukan
			1	2	3	4	5	
1.	Indikator 1 (Definisi ...)	Pernyataan 1	1	2	3	4	5	
2.		Pernyataan 2	1	2	3	4	5	
3.	Indikator 2 (Definisi ...)	Pernyataan 3	1	2	3	4	5	
		Pernyataan 4	1	2	3	4	5	

Rata-Rata Penilaian: .....

Untuk mengetahui seberapa baik alat ukur itu dalam penilaian maka nilai-nilai yang diberikan oleh penilai harus diolah lagi menjadi sebuah angka yang menunjukkan tingkat validitas alat ukur. Caranya adalah dengan merata-ratakan penilaian di setiap sub penilaian. Penilaiannya dibuat sebagai berikut:

**Tabel 4: Penilaian akhir tingkat kevalidan aspek penilaian alat ukur**

No.	Sub Penilaian	Skor	Kategori
1.	Relevansi dan representasi		
2.	Ketepatan Bahasa		
3.	Kesesuaian dengan definisi konsep dan operasional		

Kategori yang digunakan dalam penilaian adalah sebagai berikut:

**Tabel 5: Kategori penilaian tingkat kevalidan aspek-aspek penialain alat ukur**

No.	Skor	Kategori
1.	1 – 1,5	Sangat tidak valid
2.	1,6 – 2,5	Tidak valid
3.	2,6-3,5	Kurang valid
4.	3,6 – 4,0	Cukup valid
5.	4,1 – 5	Valid

Jika hasil penilaian masuk dalam kategori sangat tidak valid sampai kurang valid maka jangan digunakan alat ukur itu untuk penelitian. Alat ukur ini harus diperbaiki oleh penyusun sehingga memperoleh penilaian yang lebih baik dari penilai. Jika dalam perbaikan alat ukur sudah mendapatkan nilai cukup valid dan valid maka alat ukur itu sudah bisa digunakan penelitian. Untuk alat ukur yang masuk kategori cukup valid, sekalipun sudah bisa digunakan tetapi perlu perbaikan-perbaikan lagi jika digunakan dalam penelitian-penelitian selanjutnya.

**Daftar Pustaka**

Azwar, S. (2005). *Dasar-Dasar Psikometri*. Yogyakarta: Pustaka Pelajar.

Clark, L. A., & Watson, D. (1995). Constructing Validity: Basic Issues in Objective Scale Development. *Psychological Assessment*, 309 - 319.

Coaley, K. ( 2010). *An Introduction to Psychological Assessment and Psychometrics*. London: Sage.

Devellis, R. F. (2003). *Scale Development*. London: Sage Publications.

- Ebel, R. L. (1991). *Essentials of educational measurement (5th ed.)* , . Englewood Cliffs New Jersey: : Prentice Hall.
- Haynes, S. N., Richard, D. C., & Kubany, E. S. (1995). Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods. *Psychological Assessment*, 7, 238 - 247.
- Mesick, S. (1993). Validity. In R. Linn, *Ethnocultural Issues in the assessment* (pp. 13-104). Phoenix: Mesick, S.
- (1993) Validity, Dalam R.L. Linn(1989) *Ethnocultural Issues in the assessment* American Council on Education and Oryx Press.
- Murphy, K. R., & Davidshofer, C. O. (1991). *Psychological Testing: Principles and Applications*. New Jersey: Prentice Hall.
- Suryabrata, S. (2005). *Pengembangan Alat Ukur Psikologis*. Yogyakarta: Penerbit Andi.